

This is a repository copy of *Need, demand, supply in health care : working definitions, and their implications for defining access.*

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/179178/>

Version: Published Version

Article:

Rodriguez Santana, Idaira orcid.org/0000-0003-2022-3239, Mason, Anne Rosemary orcid.org/0000-0002-5823-3064, Gutacker, Nils orcid.org/0000-0002-2833-0621 et al. (3 more authors) (2023) Need, demand, supply in health care : working definitions, and their implications for defining access. Health Economics, Policy and Law. pp. 1-13. ISSN 1744-1331

<https://doi.org/10.1017/S1744133121000293>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:







<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

ARTICLE

Need, demand, supply in health care: working definitions, and their implications for defining access

Idaira Rodriguez Santana¹ , Anne Mason^{2*} , Nils Gutacker² , Panagiotis Kasteridis² , Rita Santos²  and Nigel Rice² 

¹HCD Economics, The Innovation Centre, Warrington, UK and ²Centre for Health Economics, University of York, York, UK

*Corresponding author. Email: anne.mason@york.ac.uk

(Received 17 December 2020; revised 5 August 2021; accepted 6 August 2021)

Abstract

Effective policymaking in health care systems begins with a clear typology of the terminology – need, demand, supply and access to care – and their interrelationships. However, the terms are contested and their meaning is rarely stated explicitly. This paper offers working definitions of need, demand and supply. We draw on the international literature and use a Venn diagram to explain the terms. We then define access to care, reviewing alternative and competing definitions from the literature. We conclude by discussing potential applications of our conceptual framework to help to understand the interrelationships and trade-offs between need, demand, supply and access in health care.

Keywords: Health services needs and demand; health services accessibility

1. Introduction

Effective policymaking and the efficient management of a health care system begin with a clear typology of the terminology – need, demand, supply and access – and their interrelationships. However, the terms are contested and their meanings are rarely stated explicitly (Williams, 1974; Culyer, 2012; Aragon *et al.*, 2017; Mason *et al.*, 2019).

Understanding the need for health care, and how this differs from the need for health that underpins it, is an important step towards quantifying current demand for health care and its future trends. It also supports decisions on the amount and type of resources the health care system requires, such as the medical and non-medical workforce and infrastructure (supply).

Informed by a review of the relevant literature, this paper offers working definitions of need, demand and supply of health care. Historically, the academic debate on concepts such as need and access has been motivated by concerns about inequities in the use of health care, e.g. differences in the consumption of health care that are not explained by differences in the need for health care (Aday and Andersen, 1974; Culyer and Wagstaff, 1993; Allin *et al.*, 2010). Like these studies, our starting point is that good health is necessary for a person to flourish (Culyer and Wagstaff, 1993) and may be viewed as a fundamental human right (UN General Assembly, 1966). However, unlike previous studies, our motivation for defining the terms is to consider broader questions about what health care to deliver, how much to deliver, where and how to deliver care, as well as for whom.

The paper is structured as follows. We use a Venn diagram to illustrate the conceptual interrelationships between the three health care elements; need, demand and supply, and the seven distinct areas created by their intersection. We then define each of the three elements, including

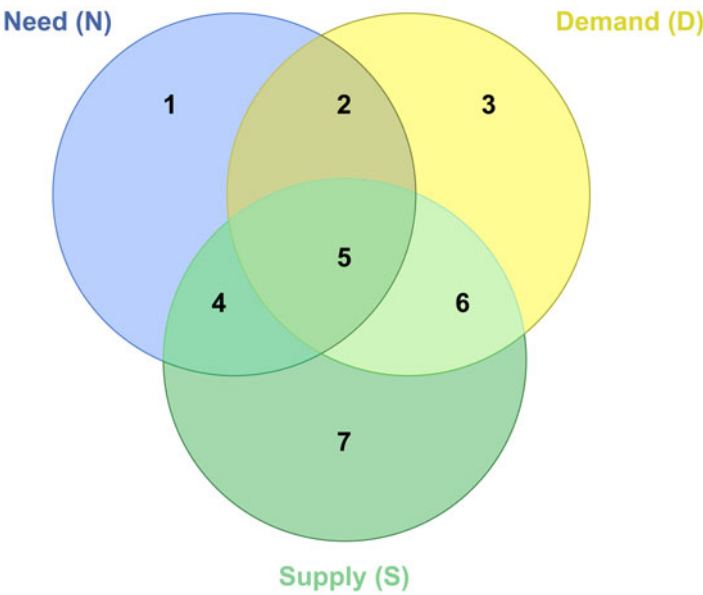


Figure 1. Venn diagram: health care need, demand and supply.
Note: Similar graphical representations have been used previously in the needs assessment literature (Stevens and Gabbay, 1991; Wright *et al.*, 1998; Stevens *et al.*, 2004) or to demonstrate the intersections of need, demand and utilisation (Hurley, 2010).

the relevant intersections, which we consider in the context of a publicly funded health care system that is mostly free at the point of use. We explain how the concept of access to care derives from the interplay between the three health care elements need, demand and supply. Finally, we consider how the conceptual framework could potentially inform policy decisions.

2. Conceptual framework and working definitions

We illustrate our conceptual framework by means of a Venn diagram (Figure 1), where health care need, demand and supply are represented by the blue, yellow and green circles, respectively. Similar graphical representations have been used previously in the needs assessment literature to illustrate drivers of and indicators for need, demand and supply (Stevens and Gabbay, 1991; Wright *et al.*, 1998; Stevens *et al.*, 2004). Moreover, the same diagram has been applied to illustrate the intersections of need, demand and utilisation (rather than supply) (Hurley, 2010). This application has a similar goal to ours, notably to highlight areas for policy attention, such as unmet need, inefficiency and inappropriate demand.

The diagram generates seven distinct areas. The size of the circles and their degree of overlap is not intended to depict the relative importance of each factor, nor to describe the current status of need, current level of demand or current provision of care within a simple health care system where care is usually free at the point of use. Nor does the uniform colour of each circle imply homogeneity in terms of types of need, the nature of demand or the care services supplied.

2.1 Need

Culyer (2012) describes the term need as “arguably the most used and least properly comprehended word in discussions of health. The meanings that attach to it are legion” (Culyer, 2012). Our working definition is derived from and builds upon the framework by Culyer and Wagstaff (1993) (Box 1).

In our conceptual framework, need is shown diagrammatically as the blue area in Figure 1 (areas 1, 2, 4 and 5). To define need, we begin with Culyer and Wagstaff’s (C&W) (1993) review

Box 1. Need – working definition

Need is the capacity to benefit from health care. Health care means treatment, prevention and supportive care that is effective – either alone or as part of a care pathway – in improving, maintaining, or slowing the deterioration of health now or in the future (or both). Need is for ‘appropriate’ health care: this excludes care that is known to be cost-ineffective and includes cost-effective care. For care of unknown cost-effectiveness, need is for the right care provided in the right place and at the right time.

of the concept of need (Culyer and Wagstaff, 1993). Table A1 (supplementary online) describes C&W’s four definitions of need, and shows step-by-step how we derived our working definition.

Like the C&W framework, we distinguish the need for health – i.e. a health deficit – from the need for health care and our definition is based on the latter. This implies that need only exists if health care exists to meet that need, i.e. need is defined by what health care is available. We also define need as need for health care that is effective (or appropriate, if its effectiveness is unknown). As in C&W, effective health care is defined in terms of improved health, or health maintenance at the point of consumption or in the future, e.g. through preventative interventions. However, our definition is broadened to encompass health care that does not directly improve health but that plays an indirect, subsidiary role by, for example, resolving diagnostic or treatment uncertainty or providing prognostic information (Newhouse, 1977).

We use the term ‘appropriate’ health care to characterise care that is either cost-effective (as in C&W) or that is of unknown cost-effectiveness but that is judged to be the right care, provided in the right setting and at the right time. The latter covers many types of care for which there is no evidence of cost-effectiveness.

Need is either met by supply (areas 4 and 5) or not (areas 1 and 2). According to our definition, if need is met then supply in areas 4 and 5 must either be directly or indirectly effective, cost-effective or ‘appropriate’ (as defined above). In contrast, unmet need, as represented by areas 1 and 2 of the Venn diagram (Figure 1), occurs when patients experience a health shortfall that could be treated in a cost-effective and appropriate manner with the available technology, but is not met by health care supply, i.e. need that does not overlap with the green supply circle (Figure 1).

Our definition has parallels with the social need taxonomy by Bradshaw (1972) who offers four definitions of need. *Normative need*, defined by professional standards, equates to areas 4 and 5 of our Venn diagram. *Expressed need* equates to areas 2 and 5. *Comparative need*, defined for individuals not in receipt of services who have similar characteristics to actual service users, is located in areas 1 and 2. Lastly, *felt need*, ‘the individual’s perceived need for services (‘want’)’ is the assumed motivation for need linked to demand (areas 2 and 5), whereas need lying within areas 1 and 4 may be either felt or not felt (unperceived). Bradshaw argues that a combination of definitions or measures is needed to identify real need and that none of the four definitions in his taxonomy is individually sufficient (Bradshaw, 1972).

Relatively few empirical studies have attempted to quantify the relationship between utilisation and unmet need. One such study is by Allin *et al.* (2010), who used Canadian Community Health Survey data to examine whether subjective assessments of unmet need could complement conventional methods of determining inequity in the utilisation of health care (Allin *et al.*, 2010). The authors defined unmet need as the difference between services received and those deemed necessary to deal appropriately with a health problem, i.e. unmet need is a shortfall of appropriate services. Allin *et al.* (2010) distinguished five types of unmet need:

- (a) **Unperceived** (by the individual) – this cannot be measured from self-reported (subjective) data but may be clinically detected, e.g. asymptomatic hypertension. As the individual is unaware of their need, they make no decision (or choice) about meeting the need. Unperceived unmet need falls into area 1.

- (b) **Chosen** (informed) – this form of unmet need results from a patient’s personal preference and is not due to a lack of awareness, education or appropriate services. For example, a patient has a condition s/he knows is treatable but decides not to seek treatment; or a patient nearing the end-of-life decides not to undergo further unpleasant treatment that would extend their life by a few months. Chosen unmet need also falls under area 1.
- (c) **Unchosen** – this covers need that is unmet, not because the individual chooses so but because of external factors outside of their control, e.g. lack of clinical staff, unaffordable travel costs, waiting lists or denial of treatment. This falls into area 2 of the Venn diagram. We refer to this type of unmet need as ‘supply-constrained’.
- (d) **Clinician validated** – the individual perceives a need for health care and demands care but does not receive care that a clinician would consider to be appropriate. The individual’s need is therefore (at least partially) unmet. Allin *et al.* (2010) contend that this dimension of utilisation is rarely addressed, with studies focusing on the level of utilisation rather than its appropriateness. This form of unmet need falls into area 2 of the Venn diagram.
- (e) **Subjective unmet expectations** – this form of unmet need is similar to clinician validated unmet need but is based on the individual’s view of the appropriateness of treatment received, rather than that of the clinician. If subjective unmet need is clinically validated, then it falls in area 2; if it is not clinically validated, then it falls into area 3 (i.e. it is not unmet need). These assignments presuppose that clinician validated need is, in some sense, genuine need.

Two of these five categories of unmet need have parallels with Bradshaw’s 1972 taxonomy of social need (Bradshaw, 1972). Bradshaw’s ‘normative’ need, based on professional value judgements, aligns with Allin’s category of ‘clinician validated’ (d). Bradshaw’s notion of ‘felt’ need, or ‘want’, mirrors Allin’s notion of ‘subjective unmet expectations’ (e).

2.2 Demand

The demand for health care is different from demand for other goods or services. First, the demand for health care is a *derived* demand arising from the fundamental demand for good health that is required for consumption and investment purposes (Grossman, 1972). Hence, health care is an input in an individual’s health production function. Other types of demand for health care can be seen as different expressions of this decision (or lack of decision) to invest in a health stock. It also offers a framework for understanding how prevention and public health – including the wider determinants of health – can influence drivers of demand for health care. Second, in many countries, patients are partially isolated from the cost of accessing care, either because of insurance coverage or because care is provided as part of a national health service (NHS). For example, in England, the NHS is generally free at the point of access although patients still face some other costs of accessing care, such as travel time or waiting times, which are important rationing factors (Blundell and Windmeijer, 2000). Third, asymmetry of information means that the doctor (agent) may demand care on behalf of the patient (principal) (Culyer, 2012). Alternatively, eligibility for care may be negotiated by the patient and the doctor (‘candidacy’) (Dixon-Woods *et al.*, 2006). In addition, patients’ ability to negotiate and to express demand may improve over time as they gain experience and understanding of their health needs and the health care system.

Our working definition of demand is informed by Culyer’s Dictionary of Health Economics (Culyer, 2012) (Box 2).

In the Venn diagram, demand is shown by the yellow circle (areas 2, 3, 5, 6).

Empirical evaluations of the drivers of the health care expenditure implicitly assume that demand can be inferred from measures of expenditure and activity. However, this captures

Box 2. Demand for health care: working definition

The **demand** for health care is the level of use at which the perceived marginal health benefits of care equal the marginal cost of accessing care. Below this point, benefits outweigh costs and individuals will continue to consume health care. The demand for health care depends on the patients' and health care professionals' perspectives of perceived benefits and costs. Benefits and costs are a function of factors such as health status, distance from providers, demographic characteristics, health literacy, etc. For example, perceived and actual health gains from care will vary with age, education and income; costs will depend on prices (e.g. co-payments for prescriptions), waiting times, time and travel costs for access, etc.

only *observed* demand, which may not reflect true demand for several reasons. Below, we describe different types of demand. These are not mutually exclusive but help to clarify the different dimensions of demand.

- (1) **Need-based** demand is demand for health care that is appropriate and hence is related to a health care need (areas 2 and 5). These are equivalent to unchosen unmet need and met need respectively. Observed utilisation of health care services includes the latter type of need but, by definition, does not include the former.
- (2) **Unnecessary** demand is depicted in area 3 and area 6 of the Venn diagram. This is demand that, by definition, is not based on need and for which care is either supplied (area 6) or not supplied (area 3). Area 3 is demand that is observable in some way but that is not based on need (according to our definition – see [Box 1](#)) and that does not lead to (further) health care utilisation. An example is a request for a GP visit motivated by a need for social interaction rather than a health need. Area 6 is demand that results in health care utilisation but is not based on need, for example, inappropriate follow-up dental or outpatient appointments.
- (3) **Avoidable** demand can arise for several reasons:
 - (a) Initially unperceived need is subsequently detected and results in demand later on in the disease pathway, e.g. an individual presents with late stage cancer.
 - (b) Some demand for health care is potentially avoidable if it arises because of behavioural risk factors, e.g. smoking, physical inactivity or substance misuse.
 - (c) Some displaced demand (see below) may also be avoidable.
 Cases of 3a and 3b are represented in area 5. Early intervention for unmet need (area 1), perhaps by proactive preventative care (area 4), could 'shift' cases out of area 5 and/or reduce the level of supply required to address them.
- (4) **Displaced** demand is demand that is displaced in time – perhaps through the lack of early intervention – or space (place). Spatial displacement refers to care in inappropriate settings, such as avoidable accident and emergency attendances or delayed discharges (area 6). In general, demand that is temporally displaced is usually avoidable, whereas spatially displaced demand can be either avoidable (e.g. patient is sent to the wrong ward by mistake) or unavoidable (e.g. patient is sent to the wrong ward because of a lack of beds on the right ward).
- (5) **Supplier-distorted demand.** Suboptimal utilisation may arise if the agent (doctor) does not convey demand on behalf of the principal (patient), such as by refusing to refer the patient for a procedure they need and request (area 2) (supplier-refused demand). The agency relationship can also lead to supplier-induced demand such as over-diagnosis or overtreatment, e.g. clinically unnecessary investigations or treatment that can result from screening programmes (area 6).

2.3 Supply

In our conceptual framework, supply is shown diagrammatically as the green area in [Figure 1](#) (areas 4, 5, 6 and 7) ([Box 3](#)).

Box 3. Supply for health care: working definition

The supply of health care includes curative and preventative services and treatments provided by the health care system, so excludes informal care and social care.

The four areas (4, 5, 6, 7) together represent the level and mix of services provided from the existing health care budget. Supply will vary over time and across geographies due to changes in the size of the budget, the production technology and the availability of scarce resources such as health care professionals.

Treatments supplied range from those that are proven to be cost-effective for a particular condition to those for which there is no evidence of effectiveness or cost-effectiveness. Whilst a treatment may or may not be effective or cost-effective for a given individual, system-level decisions will be taken based on effects averaged across relevant individuals.

Areas 4 and 5 overlap with need. According to our definition of need, these areas represent care that is appropriate. Area 6 overlaps with demand but not with need and therefore represents care that is utilised but that is not appropriate. Two types of care included within area 6 are ‘preference-sensitive care’ and ‘supply-sensitive care’. Preference-sensitive care is care that involves significant trade-offs in terms of the risks and benefits of treatment options, and where the decision to utilise care should be taken by the patient based on their preferences and values. Therefore, in some cases, the patient would not demand care if they were fully informed of the risks and benefits (Center for the Evaluative Clinical Sciences, 2007a). Supply-sensitive care arises when a health care system’s capacity constraint determines how resources are used. Supply-sensitive care has the potential to induce demand (Rosen, 2014). Induced demand is inappropriate, and so would fall in area 6. Examples include the unwarranted treatment of chronic conditions in an inpatient setting rather than in primary care, and unwarranted variations in the rates of outpatient visits, diagnostic tests or hospital beds (Center for the Evaluative Clinical Sciences, 2007b).

Area 7 is supply that is not utilised: it does not meet need or demand. There are two broad categories of care that fall into area 7. The first is ‘excess’ supply or surplus capacity. This includes deliberate ‘safety margins’ deemed appropriate because demand cannot be predicted precisely, and inefficiencies, such as duplication of services or overstocking of vaccines. The second category is supply that does not address need or demand in practice but that has the potential to do so; an example is missed appointments.

Health care differs from other commodities in part due to its supply-side characteristics: supply is relatively inelastic meaning that, in general, the volume of services and treatments cannot easily be increased in the short to medium term.

New technologies may also face barriers to entry. The effect of new technologies on the volume of care supplied depends on the additional costs and benefits over the existing technology as well as the amount of health displaced elsewhere in the system to fund the additional costs. Technological innovations for treating a previously untreatable condition have no comparator treatment. The effect of the technological innovations on supply is ambiguous, i.e. cannot be determined *a priori*. However, innovations would increase the need for health care (areas 1, 2, 4 and 5 in Figure 1): whilst the individual’s need for health existed before the new technology was invented, there was no need for health care because there was no effective health care treatment.

2.4 Access

Access is central in the performance of health care systems and its measurement has a prominent role in the health policy literature (Levesque *et al.*, 2013). As with the concept of need, there is a lack of clear

consensus regarding its meaning and measurement but most studies agree that access is a multidimensional concept (Culyer and Wagstaff, 1993; Gulliford *et al.*, 2002; McIntyre *et al.*, 2009; Levesque *et al.*, 2013). When interpreted at its broadest, it encompasses the financing, organisation and delivery of services as well as the interactions of groups and individuals that lead to utilisation (Gulliford, 2020).

We offer our working definition (Box 4), then explain our rationale. Table A2 (supplementary online) sets out five separate definitions of access reported in the literature and references the Venn diagram (Figure 1) to demonstrate how the areas map to each definition of access. Table 1 provides an overview of the five studies.

Box 4. Access to health care: working definition

Following our definition of need as being need for effective and/or appropriate health care, we define access as getting the right health care in the right place at the right time.

Our definition of access covers areas 4 and 5 of Figure 1. Below we set out our rationale for selecting this particular definition.

At the ‘macro’ level, access is about a population having the opportunity to use health care services – it is about ‘having access’ (Gulliford *et al.*, 2002) or having potential access to services (Aday and Andersen, 1974). Therefore, we might say a given population has access to health care services and this is true whether people actually use them or not. This is a supply side definition of access and so equates to availability of services (Aday and Andersen, 1974; Penchansky and Thomas, 1981; Goddard and Smith, 2001; Gulliford *et al.*, 2002; McIntyre *et al.*, 2009). In Figure 1, this definition of access is represented by areas 4–7, i.e. supply.

At the ‘micro’ level, access is about whether individuals ‘get access’ or receive care (Gulliford *et al.*, 2002). This is to do with the utilisation of services and occurs at the intersection of need and/or demand, and supply. Receipt of health care is the outcome of a series of complex processes (Dixon-Woods *et al.*, 2006). Whether or not individuals with equal need get equal access to health care services depends on the personal, financial, social and cultural barriers they face (Gulliford, 2020). This definition therefore raises questions of service accessibility (geography), affordability, acceptability, and how easy or difficult services are to navigate and what information is available. In Figure 1, this definition of access is represented by areas 4, 5 and 6.

However, we might want to know whether getting access addresses an individual’s need for health care. Utilisation may be a necessary condition for gaining access, but is it sufficient? Equating access with utilisation takes no account of the appropriateness, acceptability, effectiveness (Gulliford *et al.*, 2002) or cost-effectiveness of care – summarised by Goddard and Smith as ‘the quality of services’ (Goddard and Smith, 2001). For example, a patient who has had the wrong kidney removed has utilised health care, but has not accessed the care they need. This example would be true for all individuals. However, what constitutes quality can vary across individuals: Gulliford (2020) notes that “the needs of marginalized groups may be qualitatively different from those of the majority”, and that designing acceptable and appropriate services to address those differences is a step towards promoting equity in its vertical dimension.

In Figure 1, a definition of access that is restricted to (cost) effective health care is represented by areas 4 and 5 only: area 6 is supply that is utilised but that does not address need. This aligns with Shengelia’s concept of ‘effective coverage’ that incorporates measures of quality (broadly defined) and utilisation, conditional upon the presence of true (rather than perceived) need (Shengelia *et al.*, 2005).

This working definition represents the goal of policy. Definitions of ‘having access’ and ‘getting access’ (i.e. utilisation) can be seen as logical precursors to our working definition. In other words, if access is a process rather than an event, then deriving appropriate measures requires taking account of three factors:

Table 1. Relationship between the areas of the Venn diagram (Figure 1) and five definitions of access from the literature

Area of the Venn diagram	Is there access to health care?				
	Aday and Andersen (1974)	Penchansky and Thomas (1981)	Goddard and Smith (2001)	Gulliford <i>et al.</i> (2002)	Levesque <i>et al.</i> (2013)
1	Potential only	No	No	A: Yes B: No C: No	No
2	Potential only	No	No	A: Yes B: No C: No	No
3	Neither potential nor realised	No	No	No	No
4	Potential and realised	Yes	Yes	Yes	Yes
5	Potential and realised	Yes	Yes	Yes	Yes
6	Potential and realised	Yes	Yes	A: Yes B: Yes C: No	No
7	Potential only	No	Yes	A: Yes B: No C: No	No

Note: A: Service availability; B: Utilisation; C: Relevance and effectiveness

- (1) service availability (having access);
- (2) levels of utilisation and barriers to utilisation (getting access) that are both structural (supply side) and individual-specific (demand side); and
- (3) the effectiveness of health care supply in terms of how well it aligns with need.

An example of an overall measure of access is the effective coverage framework (Shengelia *et al.*, 2005). It can be applied at the individual level – the probability of receiving a health gain from an intervention – or at the population level (the fraction of maximum possible gain achieved) (Shengelia *et al.*, 2005).

3. Towards indicators of appropriate care

For each of the seven areas in Figure 1, Table 2 sets out the interrelationships between need, demand and supply. The implications for access are noted, and some examples offered of the types of health care situation that could be considered for measurement. We also propose some potential indicators for each area. These are not intended to be comprehensive in determining the magnitude of each area, but rather act as signals of (in)appropriate care.

4. Efficiency considerations

We also consider how the notion of efficiency fits within our conceptual framework. There are three basic types of efficiency: technical, productive and allocative efficiency (Palmer and Torgerson, 1999). Technically efficient health care arises where a given set of inputs results in the maximum possible improvement in health. This is a subset of appropriate care, so would

Table 2. Interrelationships between health care need, demand, supply and access

Area	Logic	Description of the area	Need	Unmet need	Demand	Supply	Access ^a	Example	Potential indicators
1	$(N \setminus D) \cap (N \setminus S)$ There is Need – but neither Demand nor Supply	The need for health care is unmet, and there is no access to health care. The unmet needs in this area are either (a) unperceived by the individual, or (b) needs the individual recognises but chooses not to have met.	Yes	Yes	No	No	No	(a) Unperceived unmet need Ms T is getting increasingly forgetful. Unknown to her and her family, she has dementia but been neither diagnosed nor treated. (b) Chosen unmet need Ms N has a condition she knows is treatable but she decides not to visit the GP about it.	Estimates of under-detected dementia Rate of cancer diagnoses by disease stage Acute myocardial infarction survival rates
2	$(N \cap D) \setminus S$ There is Need and Demand – but not Supply	Health care needs are expressed as demand, but supply does not meet demand. Hence, there is no access to health care. In this case, health care needs are unmet and ‘supply-constrained.’ Potential causes include access barriers, capacity constraints and waiting times.	Yes	Yes	Yes	No	No	Ms D has been depressed for some weeks and her low mood is not improving. She has asked the GP to refer her for psychotherapy. The GP has referred her but the waiting time is 6 months.	Referral waiting times
3	$(D \setminus N) \cap (D \setminus S)$ There is Demand – but neither Need nor Supply	Demand for health care is not linked to health care need and is not met by supply. There is no access to health care.	No	No	Yes	No	No	Ms Z has a painful hip. She asks her GP to refer her for an x-ray but the GP refuses her request because there is no evidence of arthritis or fracture. Ms Z has heard of a new drug for a condition she has. She asks her GP to prescribe it, but the GP refuses her request because the drug is not of proven cost-effectiveness.	Variation in referral thresholds Variation in prescribing rates
4	$(N \cap S) \setminus D$ There is Need and Supply – but not Demand	This area where need meets supply and not demand is the natural place of prevention policies. There is access to health care.	Yes	No	No	Yes	Yes	Ms G receives a letter from her child’s school saying that all pupils in her child’s class will receive a routine vaccination the following week.	Uptake of routine vaccinations and screening
5	$N \cap D \cap S$ There is Need, Demand and Supply	Health care is effective: supply meets demand (based on need) and capacity to benefit is	Yes	No	Yes	Yes	Yes	It is hay fever season and Ms W is having sinus problems. She visits her GP who prescribes a	Prescribing indicators Quality indicator performance

(Continued)

Table 2. (Continued.)

Area	Logic	Description of the area	Need	Unmet need	Demand	Supply	Access ^a	Example	Potential indicators
	Demand and Supply	positive. There is access to health care.						nasal spray and antihistamines.	emergency readmission rates delayed transfers of care
6	$(D \cap S) \setminus N$ There is Demand and Supply – but not Need	Supply meets demand but demand is not linked to health care need. There is utilisation but no access to health care.	No	No	Yes	Yes	No	<p>(a) Supplier induced demand Ms P's dental practice calls her for a dental health check-up 6 months after her last check-up. She has good dental health and, according to national guidance, only requires annual checks.</p> <p>(b) Patient induced demand Ms N has a cold. Her GP diagnoses a respiratory virus, so there is no need for antibiotics. However, in response to patient pressure he prescribes antibiotics anyway.</p> <p>Ms W is in hospital following a fall but is medically fit for discharge. The discharge team have identified a care home placement, but Ms W's family refuse to pay the fees. Ms W remains in hospital and continues to 'demand' hospital services.</p>	<p>Risk-adjusted activity rates</p> <p>Measures of observed vs expected activity</p> <p>Prescriptions for drugs of limited clinical value</p> <p>Procedures of limited clinical effectiveness</p> <p>Rates of delayed transfers of care</p>
7	$(S \setminus D) \cap (S \setminus N)$ There is Supply – but neither Need nor Demand	This case represents situations where there is excess capacity or inefficiencies in the delivery process. There is service availability but no access to health care.	No	No	No	Yes	No	In hospital F, one-third of outpatient appointments are 'did not attend' (DNA). Hospital F pays its staff the same regardless of whether or not people attend their appointments.	DNA rates

Notes: '\ ' represents the relative complement of the set, 'U' represents the union of two sets and '∩' denotes the intersection of two sets.

^aSee Access section for a detailed description of the different access definitions. Here, access implies utilisation of effective and/or appropriate health care.

fall in area 4 (which is where supply addresses need) and area 5 (where supply addresses need and demand). Note that this does not imply that all care in areas 4 and 5 is efficient: these areas may also include care produced in a technically inefficient manner.

Areas 6 and 7 are technically inefficient, as they do not address need. Therefore, the resources expended cannot improve health outcomes.

Productive efficiency occurs when health outcomes are maximised for a given cost, or when costs are minimised for a given outcome. In other words, it is closely related to cost-effectiveness. Productively efficient care must also be technically efficient, so is a subset of technically efficient care. Within our framework, productively efficient care also falls in areas 4 and 5.

Allocative efficiency is to do with allocating resources in order to maximise societal goals (Palmer and Torgerson, 1999). Allocative efficiency presupposes productive efficiency, but it also relates to questions about equity, for example, society's willingness to trade-off health losses for some groups against health gains for others. In the absence of explicit societal goals – about the level and distribution of care – it is not feasible to make definitive statements about how allocative efficiency fits within this conceptual framework. Nonetheless, the areas relevant for allocative efficiency are the same as for productive efficiency (areas 4 and 5), but unmet need (areas 1 and 2) is also pertinent: decision makers may wish to improve access to care in areas 4 and 5 for individuals, or groups of individuals, with unmet need.

5. Discussion

Our conceptual framework is a tool for understanding the types of health care that providers supply and patients demand to meet their health care needs, and also a tool for categorising different types of unmet need. The framework sheds light on the trade-offs between the elements of need, demand and supply, and informs debate about when access and utilisation occur.

In broad terms, the policy objective is to ensure demand for health care is always based on need (i.e. the demand circle in our Venn diagram lies within the need circle) (Figure 1). Need and demand may not overlap perfectly, as individuals can choose to have unmet need, even with perfect information. Supply should coincide with need-based demand, resources permitting. In other words, individuals would access and utilise services that met their needs, with the exception of chosen unmet need, and there would be no excess supply. In reality and from a system perspective, the aim is to achieve the most efficient use of limited resources of the whole system.

When reviewing the literature, we chose not to undertake a full systematic review of studies that define or analyse the concepts of need, demand, supply and access. Since our aim was to understand, synthesise and develop concepts in an area where the literature is large, diverse and complex, the value of a conventional systemic review is debatable (see Dixon-Woods *et al.*, 2006). Instead, we identified seminal papers and reviewed their bibliographies and subsequent citations akin to a pearl growing literature search. However, we do not rule out that some important studies may have been overlooked during this process.

We identify the following potential applications for our conceptual framework.

First, the proposed framework informs policy debates on the reallocation of resources and appropriate use of existing health care services. If the supply area in Figure 1 is interpreted as the maximum health care provision a health care system can offer, the opportunity costs of inappropriate use (area 6), supply (area 7) and demand for health care (area 3) become clear: the foregone benefits are those that would arise if unmet need were addressed (areas 1 and 2). Wasteful expenditure (areas 6 and 7) constitutes up to 20% of total expenditure in OECD countries (OECD and European Union, 2018). It includes missed appointments, avoidable admissions and duplication of services. Resources could be redirected towards unperceived unmet need (need of which the patient is unaware), and towards unmet need caused by informational, geographical, social, cultural or financial barriers. This is where consideration of health equity issues may also arise. Such reallocations would be expected to increase the overall allocative efficiency of the health care system.

Second, the framework reveals the potential to help to manage future demand for health care by reallocating resources to address need for health care that can improve health in the future. Reallocating resources from curative to preventive care requires careful consideration of opportunity costs. Area 4, where need and supply intersect, is an important arena for prevention policies. Activity in this area can potentially reduce future demand for effective care (area 5), but current demand pressures may make it difficult (if not impossible) to divert resources away from area 5 and into area 4.

Third, the proposed framework suggests a role for improving information for patients and the public. For example, the existence of demand that is not linked to need might be addressed by a public education campaign (area 3) or by more stringent clinical guidelines (area 6). The redirection of patients to appropriate health care services could also be facilitated through better signposting and information. Further, if individuals are more aware of the benefits of investing in their health stock then they may demand preventative care (so shifting resources from area 4 to area 5). The wider social determinants of health, i.e. factors outside of the health care system, may also be addressed by public education as well as by policy interventions.

Fourth, there is potential to explore how these considerations vary by social group or other patient characteristics such as age, gender, ethnicity, clinical condition. This could help to address inequities in both unmet and met health care needs, by allowing decision-makers to re-allocate resources to areas of greatest need in order to maximise access to health care across different socio-economic groups.

In conclusion, the proposed working definitions we offer for need, demand, supply and access are intended to provide a conceptual framework for thinking about, organising and potentially monitoring resources to ensure population health care needs are addressed in the most efficient and equitable way for a given set of resource constraints.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S1744133121000293>.

Acknowledgements. We are grateful for constructive feedback from the Department of Health and Social Care and members of our advisory group on an earlier version of this paper. All remaining errors and omissions are the responsibility of the authors.

Financial support. This study is funded by the National Institute for Health Research (NIHR) Policy Research Programme, conducted through the NIHR Policy Research Unit in Economics of Health Systems and Interface with Social Care, PR-PRU-1217-20301. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Conflict of interest. None.

References

- Aday LA and Andersen R** (1974) A framework for the study of access to medical care. *Health Services Research* **9**, 208–220.
- Allin S, Grignon M and Le Grand J** (2010) Subjective unmet need and utilization of health care services in Canada: what are the equity implications? *Social Science and Medicine* **70**, 465–472.
- Aragon MJ, Chalkley M and Goddard M** (2017) Defining and measuring unmet need to guide healthcare funding: identifying and filling the gaps. CHE RP 141. York: Centre for Health Economics, University of York, 46.
- Blundell R and Windmeijer F** (2000) Identifying demand for health resources using waiting times information. *Health Economics* **9**, 465–474.
- Bradshaw J** (1972) Taxonomy of social need. In McLachlan G (ed.), *Problems and Progress in Medical Care: Essays on Current Research, 7th Series*. Oxford: Oxford University Press, pp. 71–82.
- Center for the Evaluative Clinical Sciences** (2007a) *Preference-Sensitive Care, The Dartmouth Atlas Project*. Dartmouth, USA.
- Center for the Evaluative Clinical Sciences** (2007b) *Supply-Sensitive Care, The Dartmouth Atlas Project*. Dartmouth, USA.
- Culyer AJ** (2012) *The Dictionary of Health Economics*. Cheltenham, UK; Northampton, MA, USA: Edward Elgar Publishing Limited.
- Culyer AJ and Wagstaff A** (1993) Equity and equality in health and health care. *Journal of Health Economics* **12**, 431–457.
- Dixon-Woods M, Cavers D, Agarwal S, Annandale E, Arthur A, Harvey J, Hsu R, Katbamna S, Olsen R, Smith L, Riley R and Sutton AJ** (2006) Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Medical Research Methodology* **6**, 35.

- Goddard M and Smith P** (2001) Equity of access to health care services: theory and evidence from the UK. *Social Science & Medicine* (1982) **53**, 1149–1162.
- Grossman M** (1972) On the concept of health capital and the demand for health. *Journal of Political Economy* **80**, 223–255.
- Gulliford M** (2020) Chapter 7. Access to healthcare. In Gulliford M and Jessop E (eds), *Healthcare Public Health: Improving Health Services Through Population Science*. Oxford: Oxford University Press, pp. 69–81.
- Gulliford M, Figueroa-Munoz J, Morgan M, Hughes D, Gibson B, Beech R and Hudson M** (2002) What does 'access to health care' mean? *Journal of Health Services Research & Policy* **7**, 186–188.
- Hurley J** (2010) Chapter 8: The demand for health care. In *Health Economics*. Toronto: McGraw-Hill Ryerson, pp. 206–231.
- Levesque J-F, Harris MF and Russell G** (2013) Patient-centred access to health care: conceptualising access at the interface of health systems and populations. *International Journal for Equity in Health* **12**, 18.
- Mason A, Rodriguez Santana IDLN, Aragon Aragon MJM, Rice N, Chalkley MJ, Wittenberg R and Fernandez J-L** (2019). *Drivers of Health Care Expenditure*. York, UK: Centre for Health Economics, University of York, 1–56.
- McIntyre D, Thiede M and Birch S** (2009) Access as a policy-relevant concept in low-and middle-income countries. *Health Economics, Policy and Law* **4**, 179–193.
- Newhouse JP** (1977) Medical-care expenditure: a cross-national survey. *The Journal of Human Resources* **12**, 115–125.
- OECD and European Union** (2018) *Health at a Glance: Europe 2018: State of Health in the EU Cycle*. Paris/European Union, Brussels: OECD Publishing.
- Palmer S and Torgerson DJ** (1999) Definitions of efficiency. *BMJ* **318**, 1136.
- Penchansky R and Thomas JW** (1981) The concept of access: definition and relationship to consumer satisfaction. *Medical Care* **19**, 127–140.
- Rosen R** (2014) *Meeting Need or Fuelling Demand? Improved Access to Primary Care and Supply-Induced Demand*. London: Nuffield Trust, 18.
- Shengelia B, Tandon A, Adams OB and Murray CJ** (2005) Access, utilization, quality, and effective coverage: an integrated conceptual framework and measurement strategy. *Social Science & Medicine* (1982) **61**, 97–109.
- Stevens A and Gabbay J** (1991) Needs assessment needs assessment. *Health Trends* **23**, 20–23.
- Stevens A, Raftery J and Mant J** (2004) An introduction to HCNA. In Stevens A, Raftery J, Mant J and Simpson S (eds), *Health Care Needs Assessment. The Epidemiologically Based Needs Assessment Reviews*, vol. 1. Oxford: Radcliffe Publishing Ltd, pp. 1–6.
- UN General Assembly** (1966) International Covenant on Economic, Social and Cultural Rights. Adopted by the General Assembly of the United Nations on 16 December 1966. In Treaty Series 999 (December), Vol. 993, 171.
- Williams A** (1974). Need as a demand concept (with special reference to health). In Culyer AJ (ed.), *Economic Policies and Social Goals*. London: Martin Robertson, pp. 60–76.
- Wright J, Williams R and Wilkinson JR** (1998) Development and importance of health needs assessment. *BMJ* **316**, 1310–1313.